JOURNAL OF APPROXIMATION THEORY **14**, 51–76 (1975)

Characterization Theorems for Constrained Approximation Problems Via Optimization Theory

KENNETH R. GEHNER*

Western Electric Engineering Research Center, Princeton, New Jersey 08540

Communicated by R. C. Buck

There have been several attempts to develop a unified approach to the characterization of solutions of L^p approximation problems, for example [5] and [21]. However, the approaches developed in these papers do not readily lend themselves to handling problems where the satisfaction of additional constraints, such as interpolation or convexity conditions on the approximating function, is required. On the other hand, there have been many papers which have individually dealt with the characterization of solutions of special approximation problems with particular types of constraints, especially in the area of Chebyshev approximation. Examples of such special problems include interpolation by the approximating function [4], approximation by a monotone function [16], approximation from one side of the function to be approximated [14], [2, 6], approximation with a vector-valued norm [1, 11], simultaneous approximation of a function and its derivatives [18], and a series of papers by Taylor: [23–25].

The purpose of this paper is to develop a unified approach to the characterization of solutions of Chebyshev and L^1 approximation problems with the various types of constraints mentioned above. In addition, it is recognized that many approximation problems with non-L^p norms can easily be handled in the same manner. In Section 1 the necessary results from optimization theory are outlined. The remaining sections of the paper are devoted to applications of these results to various approximation problems: Section 2 to constrained linear Chebyshev approximation, Section 3 to rational Chebyshev approximation, Section 4 to Chebyshev approximation with a vector-valued norm, Section 5 to Chebyshev approximation with nonstandard norms, and Section 6 to constrained L^1 approximation.

1. INTRODUCTION AND PRELIMINARIES

The approach to approximation problems taken in this paper was motivated in part by the paper of Rice [20] in which a variety of approximation problems are in essence formulated as mathematical programming problems

*The author wishes to express his appreciation to Professor O. L. Mangasarian, under whose guidance a part of this research was conducted as a portion of the author's doctoral dissertation [7] in the Computer Sciences Department at the University of Wisconsin. This work was partially supported by the Mathematics Research Center at the University of Wisconsin and by National Science Foundation Contract GJ 362.

with convex constraints. Let us consider the following general optimization problem:

$$\underset{x}{\text{minimize}}\ F(x)$$

s.t.

$$\begin{aligned}
&\text{(a)}\quad G_i(x, t) \leqslant 0, \text{ for all } t \in T_i,\ i = 1,\dots, l,\\
&\text{(b)}\quad H_j(x, u) = 0, \text{ for all } u \in U_j,\ j = 1,\dots, m,\\
&\text{(c)}\quad x \in X^0,
\end{aligned} \tag{P}$$

where:

(i) each T_i and U_j is a compact subset of a complete metric space;

(ii) X^0 is an open set in R^n;

(iii) $F(x)$ and each $G_i(x, t)$ are real-valued functions which have continuous partial derivatives with respect to x for each $t \in T_i$, and each $G_i(x, t)$ is continuous in $t \in T_i$ for each $x \in X^0$; and

(iv) each $H_j(x, u)$ is a real-valued linear function in x for each $u \in U_j$, and is continuous in $u \in U_j$ for each $x \in X^0$.

The necessary and sufficient conditions for solutions of problem (P) have been previously developed [8], so the results are given here without proof.

THEOREM 1. *Let U and V be compact sets in R^n, W be an arbitrary set in R^n. Then either $uz < 0$, all $u \in U$, $vz \leqslant 0$, all $v \in V$, $wz = 0$, all $w \in W$, has a solution $z \in R^n$, or for any $u^0 \in U$, there exists $s \leqslant n$ and*

(i) *s vectors*

$$\begin{aligned}
u^i &\in U, & i &= 1,\dots, s_1,\\
v^i &\in V, & i &= s_1 + 1,\dots, s_2,\\
w^i &\in W, & i &= s_2 + 1,\dots, s,
\end{aligned}$$

(ii) *$s + 1$ real numbers λ_i, $i = 0, 1,\dots, s$ such that*

$$\lambda_i \geqslant 0 \qquad \textit{for} \quad i = 0, 1,\dots, s_2,$$

with either $\lambda_0 > 0$ or $s_1 \geqslant 1$, such that,

$$\lambda_0 u^0 + \sum_{i=1}^{s_1} \lambda_i u^i + \sum_{i=s_1+1}^{s_2} \lambda_i v^i + \sum_{i=s_2+1}^{s} \lambda_i w^i = 0,$$

but not both.

Using Theorem 1 we can prove Theorem 2.

THEOREM 2. *Let $\bar{x}$ be a local minimum of problem (P). Then there exist integers s_0 and s with $0 \leqslant s_0 \leqslant s \leqslant n$,*

(i) *there are s_0 indices i_k with $1 \leqslant i_k \leqslant l$ together with s_0 points*

$$t^k \in \hat{T}_{i_k} = \{t \in T_{i_k} \mid G_{i_k}(\bar{x}, t) = 0\}$$

for $k = 1,..., s_0$,

(ii) *there are $s - s_0$ indices j_k with $1 \leqslant j_k \leqslant m$ together with $s - s_0$ points $u^k \in U_{j_k}$ for $k = s_0 + 1,..., s$ and*

(iii) *there are $s + 1$ real numbers $\lambda_0,..., \lambda_s$ with $\lambda_0 > 0$ or $s_0 \geqslant 1$, and $\lambda_k > 0$ for $k = 1,..., s_0$, such that*

$$\text{(iv)} \quad \lambda_0 \nabla_x F(\bar{x}) + \sum_{k=1}^{s_0} \lambda_k \nabla_x G_{i_k}(\bar{x}, t^k) + \sum_{k=s_0+1}^{s} \lambda_k \nabla_x H_{j_k}(\bar{x}, u^k) = 0.$$

In order to guarantee that the conditions of Theorem 2 are meaningful, we must have $\lambda_0 > 0$. The following constraint qualifications are sufficient to prove this. As in [17], a function $F(x)$ is defined to be *pseudoconvex at* $\bar{x}$ if $\nabla F(\bar{x})(x - \bar{x}) \geqslant 0$ implies $F(x) \geqslant F(\bar{x})$.

Constraint Qualification 1 (*Modified Interior Point Condition*)

The problem (P) satisfies the *modified interior point condition* if each $G_i(x, t)$ is pseudoconvex in x for all $t \in T_i$ for $i = 1,..., l$ and there exists a point $\tilde{x} \in R^n$ which satisfies

(i) $G_i(\tilde{x}, t) < 0$, all $t \in T_i$ for $i = 1,..., l$, and

(ii) $H_j(\tilde{x}, u) = 0$, all $u \in U_j$ for $j = 1,..., m$.

Constraint Qualification 2 (*Modified Strict Inequality Condition*)

The problem (P) satisfies the *modified strict inequality condition* at a given point $\bar{x}$, where $\bar{x} \in X = \{x \in X^0 \mid G_i(x, t) \leqslant 0 \ \forall t \in T_i$ for $i = 1,..., l$ and $H_j(x, u) = 0 \ \forall u \in U_j$ for $j = 1,..., m\}$, if for any choice of integers s_0 and s with $0 \leqslant s_0 \leqslant s \leqslant n$, together with

(i) any choice of s_0 indices i_k with $1 \leqslant i_k \leqslant 1$ and s_0 points $t^k \in \hat{T}_{i_k} = \{t \in T_{i_k} \mid G_{i_k}(x, t) = 0\}$ for $k = 1,..., s_0$ and

(ii) any choice of $s - s_0$ indices j_k with $1 \leqslant j_k \leqslant m$ and $s - s_0$ points $u^k \in U_{j_k}$ for $k = s_0 + 1,..., s$, there is a vector $y = (y_1,..., y_n) \in R^n$ such that

$$\text{(iii)} \quad \sum_{q=1}^{n} y_q \nabla_{x_q} G_{i_k}(\bar{x}, t^k) < 0 \qquad \text{for} \quad k = 1,..., s_0 \quad \text{and}$$

$$\text{(iv)} \quad \sum_{q=1}^{n} y_q \nabla_{x_q} H_{j_k}(\bar{x}, u^k) = 0 \qquad \text{for} \quad k = s_0 + 1,..., s.$$

For most problems, it is usually easier to verify constraint qualification 1 rather than constraint qualification 2. Moreover, under the assumption that $G_i(x, t)$ is differentiable in x, constraint qualification 1 implies constraint qualification 2.

THOEREM 3. *Let $\bar{x}$ be a local minimum of problem (P). If either constraint qualification* 1 *or* 2 *is satisfied at $\bar{x}$, then $\lambda_0 > 0$ is guaranteed in Theorem* 2.

Under quite general convexity assumptions on the objective function and constraints of problem (P), the necessary conditions of Theorem 3 are also sufficient. Generalizing from [17], a real-valued function $G(x, t)$, where $x \in R^n$, $t \in T$ and T is an arbitrary set, is said to be *quasiconvex at $\bar{x}$* if for each x such that $G(x, t) \leqslant G(\bar{x}, t)$ $\forall t \in T$, then $G((1 - \lambda)\bar{x} + \lambda x, t) \leqslant G(\bar{x}, t)$ holds for all $0 \leqslant \lambda \leqslant 1$ for each $t \in T$. The function $G(x, t)$ is said to be *quasiconvex on a set $\Gamma \subset R^n$* if it is quasiconvex for each point $x \in \Gamma$.

THEOREM 4. *In addition to the assumptions for problem (P), let $F(x)$ be pseudoconvex on X^0, each $G_i(x, t)$ be quasiconvex on X^0, and assume that either constraint qualification* 1 *or* 2 *holds at $\bar{x}$. Then $\bar{x}$ solves problem (P) if and only if the conditions* (i)–(iv) *of Theorem* 2 *hold with $\lambda_0 > 0$.*

In the following sections of this paper we shall primarily be concerned with linear approximating functions given by $\sum_{i=1}^{n} x_i\phi_i(t)$ for all $t \in T$, where T is a compact subset of a complete metric space, and where $\{\phi_i(t)\}_{i=1}^{n}$ is a set of continuous functions on T. It shall also be assumed that $f(t)$, the function being approximated, is continuous on T. Unless explicitly stated otherwise, these assumptions hold for all the problems considered in the following sections.

2. LINEAR CHEBYSHEV APPROXIMATION

The class of problems to be considered in this section includes Chebyshev approximation problems where there are bounds on the approximation $\sum_{i=1}^{n} x_i\phi_i(t)$ and its derivatives either (i) at a certain finite number of points in the interval of approximation, or (ii) over the entire interval of approximation. The general problem can be written as

$$\underset{x,\tau}{\text{minimize}}\ \tau$$

s.t.

$$\text{(i)}\quad -\tau \leqslant \sum_{i=1}^{n} x_i\phi_i(t) - f(t) \leqslant \tau, \qquad \text{all } t \in T, \tag{1}$$

$$\text{(ii)}\quad l_k(t) \leqslant \sum_{i=1}^{n} x_i \phi_i^{(j_k)}(t) \leqslant u_k(t), \qquad \text{all } t \in T, \quad \text{for} \quad k = 1,..., K_0$$

$$\text{(iii)}\quad \gamma_{1k} \leqslant \sum_{i=1}^{n} x_i \phi_i^{(j_k)}(\bar{t}^k) \leqslant \gamma_{2k}, \qquad \text{for} \quad k = K_0 + 1,..., K_1, \tag{1}$$

$$\text{(iv)}\quad \sum_{i=1}^{n} x_i \phi_i^{(j_k)}(\bar{t}^k) = \gamma_k \qquad \text{for} \quad k = K_1 + 1,..., K,$$

where the indices j_k are prescribed nonnegative integers, for each $k = 1,..., K_0$, $l_k(t) \leqslant u_k(t)$ $\forall t \in T$ with both $l_k(t)$ and $u_k(t)$ being continuous functions on T, each $\bar{t}^k \in T$, and $\gamma_{1k} < \gamma_{2k}$ for $k = K_0 + 1,..., K_1$.

Many problems previously considered in the approximation theory literature are special cases of (1). For instance, first consider the case where $K_0 = 1$, $K = K_0$ (no constraints (iii) and (iv)), and $j_1 = 0$, which results in the problem:

$$\underset{x,\tau}{\text{minimize}}\ \tau$$

s.t.

$$\text{(i)}\quad -\tau \leqslant \sum_{i=1}^{n} x_i \phi_i(t) - f(t) \leqslant \tau, \tag{2}$$

$$\text{(ii)}\quad l(t) \leqslant \sum_{i=1}^{n} x_i \phi_i(t) \leqslant u(t),$$

all $t \in T$. If we let $u(t) \equiv M$ for all $t \in T$, where M is a very large positive number, and $l(t) \equiv 0$ for all $t \in T$, then we have the problem of nonnegative approximation which has been studied by Jones and Karlovitz [12] under the condition that $\{\phi_i(t)\}_{i=1}^{n}$ forms a Haar set on T. If we let $l(t) \equiv -M$ $\forall t \in T$, where M is a large positive number, and $u(t) = f(t)$, then we have the problem of one-sided approximation (approximation from below), which has been studied by Kammerer [13]. Generalized versions of problem (2) have been studied by Taylor and others in a series of papers: Taylor [23]–[25], Taylor and Schumaker [22], and Taylor and Winter [26].

A second class of problems can be brought into consideration by setting $K = K_0$ (no constraints (iii) and (iv) in (1)) and either (i) $l_k(t) \equiv 0$ and $u_k(t) \equiv M$, or (ii) $l_k(t) \equiv -M$ and $u_k(t) \equiv 0$ where M is a very large positive number in either case. This leads to the problem

$$\underset{x,\tau}{\text{minimize}}\ \tau$$

s.t.

$$\text{(i)}\quad -\tau \leqslant \sum_{i=1}^{n} x_i \phi_i(t) - f(t) \leqslant \tau, \tag{3}$$

(ii) $\epsilon_k \left[\sum_{i=1}^{n} x_i \phi^{(j_k)}(t) \right] \leqslant 0, \qquad$ for $k = 1,\ldots, K$ and $\epsilon_k = \pm 1$,

all $t \in T$. Lorentz and Zeller [16] have considered a special case of (3), where $T = [a, b]$, an interval of the real line, the set $\{\phi_i(t)\}_{i=1}^n = \{t^{i-1}\}_{i=1}^n$ (polynomial approximation), and $K \leqslant n$. Special cases of interest for (3) are

(i) $K = 1$ and $j_1 = 1$: approximation by a monotone function,

(ii) $K = 1$ and $j_1 = 2$: approximation by a convex function.

The main result of this section is the following characterization theorem for the original problem (1).

THEOREM 5. *Assume that $f(t)$ is not in the span of $\{\phi_i(t)\}_{i=1}^n$ and that either constraint qualification 1 or 2 holds for problem (1). Then a point $(\tau^*, x_1^*,\ldots, x_n^*)$ which is feasible for (1) yields an optimal approximation if and only if the origin of R^n can be written as a convex combination of at most $n + 1$ points from the union of the sets*

$$X_\tau = \left\{ e(t) \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix} \middle| \, | e(t) | = \| e \|_T = \tau^* \right\},$$

where $e(t) = \sum_{i=1}^n x_i^* \phi_i(t) - f(t)$ and $\| e \| = \max_{t \in T} | e(t) |$,

$$X_{l_k} = \left\{ - \begin{pmatrix} \phi_1^{(j_k)}(t) \\ \vdots \\ \phi_n^{(j_k)}(t) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i^{(j_k)}(t) = l_k(t) \right\},$$

$$X_{u_k} = \left\{ + \begin{pmatrix} \phi_1^{(j_k)}(t) \\ \vdots \\ \phi_n^{(j_k)}(t) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i^{(j_k)}(t) = u_k(t) \right\},$$

$$X_{1k} = \left\{ - \begin{pmatrix} \phi_1^{(j_k)}(\bar{t}^k) \\ \vdots \\ \phi_n^{(j_k)}(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i^{(j_k)}(\bar{t}^k) = \gamma_{1k} \right\},$$

$$X_{2k} = \left\{ + \begin{pmatrix} \phi_1^{(j_k)}(\bar{t}^k) \\ \vdots \\ \phi_n^{(j_k)}(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i^{(j_k)}(\bar{t}^k) = \gamma_{2k} \right\},$$

$$X_k = \left\{ \pm \begin{pmatrix} \phi_1^{(j_k)}(\bar{t}^k) \\ \vdots \\ \phi_n^{(j_k)}(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i^{(j_k)}(\bar{t}^k) = \gamma_k \right\},$$

with at least one point from the set X_τ included nontrivially.

Proof. Since each of the constraints (i)–(iv) are linear in x and τ and either constraint qualification 1 or 2 holds, the characterization conditions of Theorem 4 hold. Thus, there are integers s_0, s_1, s_2, and s with $0 \leqslant s_0 \leqslant s_1 \leqslant s_2 \leqslant s \leqslant n$ together with $t^q \in T$ for $q = 1,..., s_1$, $1 \leqslant k_q \leqslant K_0$ for $q = s_0 + 1,..., s_1$, $K_0 + 1 \leqslant k_q \leqslant K_1$ for $q = s_1 + 1,..., s_2$, and $K_1 + 1 \leqslant k_q \leqslant K_2$ for $q = s_2 + 1,..., s$ such that

$$\begin{aligned}
&\bar{\lambda}_0 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \sum_{q=1}^{s_0} \bar{\lambda}_q \begin{bmatrix} -1 \\ (-1)^{\epsilon_q} \phi_1(t^q) \\ \vdots \\ (-1)^{\epsilon_q} \phi_n(t^q) \end{bmatrix} + \sum_{q=s_0+1}^{s_1} \bar{\lambda}_q \begin{bmatrix} 0 \\ (-1)^{\epsilon_q} \phi_1^{(j_{k_q})}(t^q) \\ \vdots \\ (-1)^{\epsilon_q} \phi_n^{(j_{k_q})}(t^q) \end{bmatrix} \\
&+ \sum_{q=s_1+1}^{s_2} \bar{\lambda}_q \begin{bmatrix} 0 \\ (-1)^{\epsilon_q} \phi_1^{(j_{k_q})}(\bar{t}^{k_q}) \\ \vdots \\ (-1)^{\epsilon_q} \phi_n^{(j_{k_q})}(\bar{t}^{k_q}) \end{bmatrix} + \sum_{q=s_2+1}^{s} \bar{\lambda}_q \begin{bmatrix} 0 \\ (-1)^{\epsilon_q} \phi_1^{(j_{k_q})}(\bar{t}^{k_q}) \\ \vdots \\ (-1)^{\epsilon_q} \phi_n^{(j_{k_q})}(\bar{t}^{k_q}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\end{aligned} \tag{4}$$

where $\bar{\lambda}_q > 0$ for $q = 0, 1,..., s$ and $\epsilon_q = 0$ or 1 such that

(i) for $q = 1,..., s_0$,

$$\epsilon_q = \begin{cases} 0 \text{ if } f(t^q) - \sum_{i=1}^{n} x_i^* \phi_i(t^q) = -\tau^*, \\ 1 \text{ if } f(t^q) - \sum_{i=1}^{n} x_i^* \phi_i(t^q) = +\tau^*, \end{cases}$$

(ii) for $q = s_0 + 1,..., s_1$,

$$\epsilon_q = \begin{cases} 0 \text{ if } \sum_{i=1}^{n} x^* \phi_i^{(j_{k_q})}(t^q) = u_{k_q}(t^q), \\ 1 \text{ if } \sum_{i=1}^{n} x_i^* \phi^{(j_{k_q})}(t^q) = l_{k_q}(t^q). \end{cases}$$

(iii) for $q = s_1 + 1,..., s_2$,

$$\epsilon_q = \begin{cases} 0 \text{ if } \sum_{i=1}^{n} x_i^* \phi_i^{(j_{k_q})}(\bar{t}^{k_q}) = \gamma_{2k_q}, \\ 1 \text{ if } \sum_{i=1}^{n} x_i^* \phi_i^{(j_{k_q})}(\bar{t}^{k_q}) = \gamma_{1k_q}, \end{cases}$$

and

(iv) for $q = s_2 + 1, \ldots, s$, the ϵ_q is chosen appropriately so as to force $\bar{\lambda}_q > 0$.

Since $f(t)$ is not in the span of $\{\phi_i(t)\}_{i=1}^n$, $\tau^* > 0$. Thus, we have $\bar{\lambda}_q = (\bar{\lambda}_q/\tau^*) \cdot \tau^*$ for $q = 1, \ldots, s_0$ and

$$e(t^q) = \sum_{i=1}^n x_i^* \phi_i(t^q) - f(t^q) = \begin{cases} \tau^* & \text{if } \epsilon_q = 0, \\ -\tau^* & \text{if } \epsilon_q = 1. \end{cases}$$

Noting that $\bar{\lambda}_0 > 0$ implies that $s_0 \geqslant 1$, i.e., there is at least one vector from X_τ in the linear combination, substituting the previous observation into (4), and normalizing the resulting coefficients such that they sum to 1, it is shown that the origin of R^n can be written as a convex combination of vectors from the specified sets. Since the steps above are reversible to obtain (4), the convex combination condition is also sufficient. Q.E.D.

It must be noted that in order to apply Theorem 4 either constraint qualification 1 or 2 must be satisfied. For an example of an approximation problem which does not satisfy the constraint qualifications and which consequently does not conform to the results of Theorem 4, see [8]. However, in many important cases these constraint qualifications are satisfied. For example, for ordinary Chebyshev approximation or Chebyshev approximation with interpolation (see [8]) constraint qualification 1 is immediately satisfied since τ can be made large enough in constraint (i) so that the inequalities are strict over all of T. In other problems this can be shown rather easily, for example in the generalization (a function $g_k(t)$ replaces 0) of the Lorentz and Zeller problem [16]:

$$\begin{aligned} &\min_{\tau, x} \tau \\ &\text{s.t.} \\ &\text{(i)} \quad -\tau \leqslant \sum_{i=1}^n x_i t^{i-1} - f(t) \leqslant \tau, \\ &\text{(ii)} \quad \epsilon_k \sum_{i=j_k}^n [j_k(j_k - 1) \cdots 1 \cdot x_i t^{i-j_k-1}] \leqslant g_k(t) \qquad \text{for} \quad k = 1, \ldots, K, \end{aligned} \tag{5}$$

$\epsilon_k = \pm 1$, for all $t \in T$ (not necessarily an interval) there is a polynomial $P(\tilde{x}, t) = \sum_{i=1}^n \tilde{x}_i t^{i-1}$ and $\tilde{\tau}$ such that constraints (i) and (ii) of (5) are satisfied strictly. Indeed, since T is a compact subset of the real line, $T \subset [a, b]$ for some interval $[a, b]$, and thus by successively choosing $A_K, A_{K-1}, \ldots, A_1$ (assuming $j_1 < \cdots < j_K$) in

$$P(t) = A_K(t - a)^{j_K} + A_{K-1}(t - a)^{j_{K-1}} \cdots + A_1(t - a)^{j_1},$$

so that $P(t)$ satisfies constraint (ii) strictly, we have the required $P(\tilde{x}, t) = P(t)$.

Finally, under additional conditions it is possible to prove a generalization of the classical alternation theorem (see [3]) which yields an interesting interpretation of the optimality conditions. For the purpose of this characterization, notice that each of the constraints (i)–(iv) in (1) can be written as a pair of inequalities bounding either $\sum_{i=1}^{n} x_i\phi_i(t)$ or a derivative of $\sum_{i=1}^{n} x_i\phi_i(t)$ from above and below, and let us use the term *upper bounding constraint at* $t \in T$ when the upper bound on any of these two sided inequalities is active, i.e., an equality, at t and the term *lower bounding constraint at* $t \in T$ when the lower bound on any of the inequalities is active at t.

THEOREM 6. *Assume that T is a compact subset of the real line, $\{\phi_i(t)\}_{i=1}^{n}$ is a Haar set on T, $f(t)$ is not in the span of $\{\phi_i(t)\}_{i=1}^{n}$, no derivatives are involved in the constraints* (ii)–(iv), *and that either constraint qualification* 1 *or* 2 *holds for problem* (1). *Then a point $(\tau^*, x_1^*,..., x_n^*)$ which is feasible for* (1) *yields an optimal approximation if and only if there exist $n+1$ points t^{k_q} for $q = 1,..., n+1$ with either $t^{k_q} \in T$ or $t^{k_q} = \bar{t}^k$ for some $k = 1,..., K$ such that at least one t^{k_q} forces one of the constraints* (i)–(iv) *to be active and with $t^{k_1} < \cdots < t^{k_{n+1}}$ the active constraints alternate from an upper bounding constraint to a lower bounding constraint at consecutive t^{k_q}.*

Proof. The theorem follows at once from Theorem 5 and the characterization lemma for the origin to be in the convex hull of a Haar system (see [3, p. 74]), which forces the signs of the vectors evalauted at consecutive points to alternate. Thus, the active constraints must alternate from upper bounding to lower bounding constraints in problem (1). Q.E.D.

3. RATIONAL CHEBYSHEV APPROXIMATION

Given two sets $\{\phi_i(t)\}_{i=1}^{n}$ and $\{\psi_j(t)\}_{j=1}^{m}$ of continuous real-valued functions for all $t \in T$, where T is a compact metric space, and a continuous real-valued function $f(t)$ on T, the problem of generalized rational approximation is to find parameters $x_1^*,..., x_n^*$ and $y_1^*,..., y_m^*$ such that

$$\underset{t\in T}{\text{maximum}} \left| \frac{\sum_{i=1}^{n} x_i^*\phi_i(t)}{\sum_{j=1}^{m} y_j^*\Psi_j(t)} - f(t) \right| = \underset{\substack{x_1,\ldots,x_n\\ y_1,\ldots,y_m}}{\text{minimum}} \max_{t\in T} \left| \frac{\sum_{i=1}^{n} x_i\phi_i(t)}{\sum_{j=1}^{m} y_j\Psi_j(t)} - f(t) \right| .$$

This problem can be stated in a more convenient form as:

$$\underset{\substack{\tau,x_1,\ldots,x_n,\\ y_1,\ldots,y_m}}{\text{minimize}}\ \tau$$

s.t.

$$\begin{aligned}
&\text{(i)} \quad \sum_{i=1}^{n} x_i\phi_i(t) - (\tau + f(t)) \sum_{j=1}^{m} y_j\Psi_j(t) \leqslant 0, \qquad \text{all} \quad t \in T, \\
&\text{(ii)} \quad -\sum_{i=1}^{n} x_i\phi_i(t) - (\tau - f(t)) \sum_{j=1}^{m} y_j\Psi_j(t) \leqslant 0, \qquad \text{all} \quad t \in T, \\
&\text{(iii)} \quad \sum_{i=1}^{m} y_j\Psi_j(t) > 0, \qquad \text{all} \quad t \in T.
\end{aligned} \tag{6}$$

The first observation is that the set of variables $(\tau, x_1, ..., x_n, y_1, ..., y_m)$ which satisfies the relationship (iii) is an open set X^0 in R^{n+m+1}, which follows from the continuity of the functions $\{\psi_j(t)\}_{j=1}^{m}$. Next, observe that neither constraint (i) or (ii) is necessarily pseudo or quasiconvex in the parameters $(\tau, x_1, ..., x_n, y_1, ..., y_m)$ for all values of these parameters. Thus, because the constraints do not satisfy the appropriate conditions, neither the necessary condition Theorem 3 nor the characterization Theorem 4 can be applied to this problem. However, the following result can still be proved.

THEOREM 7. *A point $(\tau^*, x_1^*, ..., x_n^*, y_1^*, ..., y_m^*)$ solves problem* (6) *if and only if there are s points $t^k \in \{t \in T \mid | R^*(t) - f(t)| = \tau^*\}$ and s real numbers γ_k with each $\gamma_k \neq 0$ and $R^*(t^k) - f(t^k) = (\operatorname{sgn} \gamma_k) \| R^*(t) - f(t)\|_T$, where $1 \leqslant s \leqslant n + m + 1$, such that*

$$\sum_{k=1}^{s} \gamma_k \begin{pmatrix} \phi_1(t^k) \\ \vdots \\ \phi_n(t^k) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{7}$$

and

$$\sum_{k=1}^{s} \gamma_k R^*(t^k) \begin{pmatrix} \psi_1(t^k) \\ \vdots \\ \psi_m(t^k) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{8}$$

where

$$R^*(t) = \sum_{i=1}^{n} x_i^*\phi_i(t) \Big/ \sum_{i=1}^{m} y_j^*\psi_j(t)$$

and

$$\| R^*(t) - f(t)\|_T = \max_{t \in T} | R^*(t) - f(t)| = \tau^*.$$

Proof. If $(\tau^*, x_1^*, ..., x_n^*, y_1^*, ..., y_m^*)$ solves problem (6), then by Theorem 2 there exist real numbers $\lambda_0 \geqslant 0$, $\lambda_k > 0$ for $k = 1, ..., s$ with

$1 \leqslant s \leqslant m + n + 1$ such that

$$\lambda_0 \begin{bmatrix} 1 \\ \hline 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix} + \sum_{k=1}^{s} \lambda_k \begin{bmatrix} -\sum_{j=1}^{m} y_j{}^* \Psi_j(t^k) \\ \hline (-1)^{\epsilon_k} \phi_1(t^k) \\ \vdots \\ (-1)^{\epsilon_k} \phi_n(t^k) \\ \hline -(\tau^* + (-1)^{\epsilon_k} f(t^k)) \Psi_1(t^k) \\ \vdots \\ -(\tau^* + (-1)^{\epsilon_k} f(t^k)) \Psi_m(t^k) \end{bmatrix} = \begin{bmatrix} 0 \\ \hline 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{9}$$

where $\epsilon_k = 0$ if constraint (i) is active at t^k and $\epsilon_k = 1$ if constraint (ii) is active at t^k. From the first component of (9) it follows that

$$\sum_{k=1}^{s} \lambda_k \left[\sum_{j=1}^{m} \gamma_j{}^* \psi_j(t^k) \right] = \lambda_0 \text{ and so } \lambda_0 > 0$$

must hold by constraint (iii) and the fact that $\lambda_k > 0$ for $k = 1,..., s$. Since $(\tau^* + (-1)^{\epsilon_k} f(t^k)) = (-1)^{\epsilon_k} R^*(t^k)$, $(-1)^{\epsilon_k}[R^*(t^k) - f(t^k)] = \tau^*$ and

$$\sum_{k=1}^{s} \lambda_k (-1)^{\epsilon_k} R^*(t^k) \begin{pmatrix} \psi_1(t^k) \\ \vdots \\ \psi_m(t^k) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

from the last set of m components of (9). Thus, by defining $\gamma_k = (-1)^{\epsilon_k} \lambda_k$, the necessary conditions (7) and (8) follow.

The sufficiency of conditions (7) and (8) is easily proved as follows. Suppose (7) and (8) hold. Then if the function $R^*(t)$ is not a best approximation to $f(t)$, there is an $R(t) = P(t)/Q(t)$ such that $\| R(t) - f(t)\|_T < \| R^*(t) - f(t)\|_T = \tau^*$ where $P(t) = \sum_{i=1} x_i \phi_i(t)$ and $Q(t) = \sum_{j=1} y_j \psi_j(t) > 0$ for all $t \in T$. Furthermore, we have $(\operatorname{sgn} \gamma_k)(R(t^k) - f(t^k)) \leqslant \| R(t) - f(t)\|_T < \| R^*(t) - f(t)\|_T = (\operatorname{sgn} \gamma_k)(R^*(t^k) - f(t^k))$ so $(\operatorname{sgn} \gamma_k)(R^*(t^k) - R(t^k) > 0$ for $k = 1,..., s$. But since $Q(t) > 0$ for all $t \in T$, it follows that

$$(\operatorname{sgn} \gamma_k)(R^*(t^k)\, Q(t^k) - P(t^k)) > 0 \qquad \text{for} \quad k = 1,..., s. \tag{10}$$

Multiplying (7) by $(x_1, ..., x_n)$ and subtracting the results from (8) multiplied by $(y_1, ..., y_m)$ it follows that

$$\sum_{k=1}^{s} \gamma_k [R^*(t^k)\, Q(t^k) - P(t^k)] = 0,$$

which contradicts (10). Q.E.D.

Although Theorem 7 is well known [3, p. 160], the merit of this approach to the theorem is that it is directly generalizable to problems of generalized rational approximation with auxillary constraints such as interpolation conditions, which have not been extensively studied. For the problem of rational approximation with additional interpolation requirements the following characterization theorem can be obtained.

THEOREM 8. *Consider problem* (6) *with the additional constraint,*

$$\text{(iv)} \quad \sum_{i=1}^{n} x_i\phi_i(t) - b(t)\sum_{j=1}^{m} y_j\psi_j(t) = 0, \qquad \textit{for all} \quad t \in S,$$

where S is an arbitrary set. Then a point $(\tau^*, x_1^*,..., x_n^*, y_1^*,..., y_m^*)$ solves problem (6) *with constraint* (iv) *if and only if there are integers s_0 and s with $1 \leqslant s_0 \leqslant s \leqslant n + m + 1$ with the properties that*

(i) *there are s_0 points $t^k \in \{t \in T \mid | R^*(t) - f(t)| = \| R^*(t) - f(t)\|_T\}$ and s_0 real numbers γ_k with each $\gamma_k \neq 0$ and*

$$R^*(t^k) - f(t^k) = (\operatorname{sgn} \gamma_k) \| R^*(t) - f(t)\|_T \qquad \textit{for} \quad k = 1,..., s_0\,,$$

(ii) *there are $s - s_0$ points $t^k \in S$ and $s - s_0$ real numbers $\gamma_k \neq 0$ such that*

$$\sum_{k=1}^{s} \gamma_k \begin{pmatrix} \phi_1(t^k) \\ \vdots \\ \phi(t^k) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{11}$$

$$\sum_{k=1}^{s_0} \gamma_k R^*(t^k)\begin{pmatrix} \psi_1(t^k) \\ \vdots \\ \psi_m(t^k) \end{pmatrix} + \sum_{k=s_0+1}^{s} \gamma_k b(t^k)\begin{pmatrix} \psi_1(t^k) \\ \vdots \\ \psi_m(t^k) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{12}$$

Proof. The necessity follows as in Theorem 7 since from Theorem 2 either $\lambda_0 > 0$ or $s_0 \geqslant 1$ which implies $\lambda_0 > 0$ because $\sum_{k=1}^{s_0} \lambda_k[\sum_{j=1}^{m} y_j\psi_j(t^k)] = \lambda_0$ and constraint (iii) holds. For the sufficiency, assume (11) and (12) hold. Also assume that there is a rational function $R(t) = P(t)/Q(t)$ where $P(t) = \sum_{i=1} x_i\phi_i(t)$ and $Q(t) = \sum_{i=1}^{n} y_i\psi_i(t)$ with $P(t) = b(t)\,Q(t)$ for all $t \in S$ such that $R(t)$ is a better approximation to $f(t)$ than $R^*(t)$. Then as in the proof of Theorem 7,

$$(\operatorname{sgn} \gamma_k)(R^*(t^k) - R(t^k)) > 0 \qquad \text{for} \quad k = 1,..., s_0\,. \tag{13}$$

Multiplying (11) by $(x_1,..., x_n)$ and subtracting the result from (12) multiplied by $(y_1,..., y_m)$,

$$\sum_{k=1}^{s_0} \gamma_k[R^*(t^k)\,Q(t^k) - P(t^k)] + \sum_{k=s_0+1}^{s} \gamma_k[b(t^k)\,Q(t^k) - P(t^k)] = 0.$$

But since $P(t) = b(t)\, Q(t)$ for all $t \in S$, this becomes

$$\sum_{k=1}^{s_0} \gamma_k [R^*(t^k)\, Q(t^k) - P(t^k)] = 0,$$

which contradicts (13). Q.E.D.

4. Chebyshev Approximation with a Vector-Valued Norm

In this section, problems of the following form shall be considered. Given a vector-valued objective function $[F_1(x),\ldots, F_Q(x)]$, problem (P_v) is given by

$$\text{"minimize"}\ [F_1(x),\ldots, F_Q(x)]$$

s.t.

$$\begin{aligned} &\text{(a)}\quad G_i(x, t) \leqslant 0, \text{ all } t \in T_i \text{ for } i = 1,\ldots, l, \\ &\text{(b)}\quad H_j(x, u) = 0, \text{ all } u \in U_j \text{ for } j = 1,\ldots, m, \\ &\text{(c)}\quad x \in X^0, \end{aligned} \qquad (P_v)$$

where:

(i) each T_i and U_j is a compact set of a complete metric space;

(ii) X^0 is an open set in R^n;

(iii) each $F_q(x)$ is a convex function in x which has continuous partial derivatives with respect to x; and

(iv) each $G_i(x, t)(H_j(x, u))$ is a quasiconvex (linear) function in x which has continuous partial derivatives with respect to x for each $t \in T_i$ $(u \in U_j)$.

A feasible point $\bar{x}$ is said to be *efficient* for problem (P_v) if there does not exist a point $\tilde{x}$ which is feasible for problem (P_v) such that $[F_1(\tilde{x}),\ldots, F_Q(\tilde{x})] \leqslant [F_1(\bar{x}),\ldots, F_Q(\bar{x})]$ and $F_q(\tilde{x}) < F_q(\bar{x})$ for at least one $q = 1,\ldots, Q$. In other words, $\bar{x}$ is efficient for (P_v) if no improvement can be made in any component of the objective function without sacrificing in another component. Furthermore (P_v) is said satisfy the *vector constraint qualification* 1 at $\bar{x}$ if for each $q_0 = 1,\ldots, Q$ there is a point $\tilde{x} \in R^n$ which satisfies

(i) $F_q(\tilde{x}) \leqslant F_q(\bar{x})$, $q = 1,\ldots, Q$, $q \neq q_0$,

(ii) $G_i(\tilde{x}, t) < 0$, all $t \in T_i$ for $i = 1,\ldots, l$,

(iii) $H_j(\tilde{x}, u) = 0$, all $u \in U_j$ for $j = 1,\ldots, m$.

Similarly, problem (P_v) is said to satisfy the *vector constraint qualification* 2 at $\bar{x}$ if for each $q_0 = 1,\ldots, Q$ and any choice of integers $0 \leqslant s_0 \leqslant s_1 \leqslant s \leqslant n$ together with

(i) any choice of s_0 indices q_k with $1 \leqslant q_k \leqslant Q$ and $q_k \neq q_0$ for $k = 1,\ldots, s_0$,

(ii) any choice of $s_1 - s_0$ indices i_k with $1 \leqslant i_k \leqslant l$ and $s_1 - s_0$ points $t^k \in \hat{T}_{i_k} = \{t \in T_{i_k} \mid G_{i_k}(\bar{x}, t) = 0\}$ for $k = s_0 + 1,\ldots, s_1$,

(iii) any choice of $s - s_1$ indices j_k with $1 \leqslant j_k \leqslant m$ and $s - s_1$ points $u^k \in U_{j_k}$ for $k = s_1 + 1,\ldots, s$, there exists a vector $y \in R^n$ such that

(iv) $\nabla_x F_{q_k}(\bar{x})\, y \leqslant 0$, $k = 1,\ldots, s_0$,

(v) $\nabla_x G_{i_k}(\bar{x}, t^k)\, y < 0$, $k = s_0 + 1,\ldots, s_1$, and

(vi) $\nabla_x H_{j_k}(\bar{x}, u^k)\, y = 0$, $k = s_1 + 1,\ldots, s$.

Note that constraint qualification 1 implies qualification 2.

The result which permits useful analysis for problem (P_v) is as follows.

THEOREM 9. *Assume that (P_v) satisfies vector constraint qualification 1 or 2. A point $\bar{x}$ is efficient for problem (P_v) if and only if $\bar{x}$ solves problem (P_α) where*

$$\text{minimize} \sum_{q=1}^{Q} \alpha_q F_q(x)$$

s.t.

(i) $G_i(x, t) \leqslant 0$, *all* $t \in T_i$ *for* $i = 1,\ldots, l$,

(ii) $H_j(x, u) = 0$, *all* $u \in U_j$ *for* $j = 1,\ldots, m$,

(iii) $x \in X^0$, (P_α)

for some $\alpha \in R^Q$ *with each* $\alpha_q > 0$.

Proof. If $\bar{x}$ is efficient for (P_v), then for each $q_0 = 1,\ldots, Q$, there is no solution $z \in R^n$ to the system

$$\begin{aligned} \nabla F_{q_0}(\bar{x})\, z &< 0, \\ \nabla G_i(\bar{x}, t)\, z &< 0, \quad \text{all } t \in T_i \text{ for } i = 1,\ldots, l, \\ \nabla F_q(\bar{x})\, z &\leqslant 0, \quad q = 1,\ldots, Q,\ q \neq q_0, \\ \nabla H_j(\bar{x}, u)\, z &= 0, \quad \text{all } u \in U_j \text{ for } j = 1,\ldots, m. \end{aligned} \tag{14}$$

The proof of this is very similar to the proof in [10], so the details are not given here. Since (14) has no solution, by Theorem 1,

$$\lambda_0 \nabla F_{q_0}(\bar{x}) + \sum_{k=1}^{s_0} \lambda_k \nabla G_{i_k}(\bar{x}, t^k) + \sum_{k=s_0+1}^{s_1} \lambda_k \nabla F_{q_k}(\bar{x}) + \sum_{k=s_1+1}^{s} \lambda_k \nabla H_{j_k}(\bar{x}, u^k) = 0, \tag{15}$$

for some $\lambda_0 \geqslant 0$, $\lambda_k > 0$ for $k = 1,..., s_0$, $\lambda_k \geqslant 0$ for $k = s_0 + 1,..., s_1$ with either $\lambda_0 > 0$ or $s_0 \geqslant 1$. If $\lambda_0 = 0$, then by vector constraint qualification 2, there exists a $y \in R^n$ such that

$$\underbrace{\sum_{k=1}^{s_0} \lambda_k \nabla G_{i_k}(\bar{x}, t^k)\, y}_{\leftarrow\, <0\, \rightarrow} + \underbrace{\sum_{k=s_0+1}^{s_1} \lambda_k \nabla F_{q_k}(\bar{x})\, y}_{\leftarrow\, \leqslant 0\, \rightarrow} + \underbrace{\sum_{k=s_1+1}^{s} \lambda_k \nabla H_{j_k}(\bar{x}, u^k)\, y}_{\leftarrow\, =0\, \rightarrow} = 0.$$

But this is a contradiction so $\lambda_0 > 0$ holds. Since (15) holds for each $q_0 = 1,..., Q$, by summing these equations it follows that

$$\nabla \left[\sum_{q=1}^{Q} \alpha_q F_q(\bar{x}) \right] + \sum_{k=1}^{s_0} \lambda_k \nabla G_{i_k}(\bar{x}, t^k) + \sum_{k=s_0+1}^{s} \lambda_k \nabla H_{j_k}(\bar{x}, u^k) = 0, \tag{16}$$

with each $\alpha_q > 0$ and each $\lambda_k > 0$ for $k = 1,..., s_0$. By Theorem 4, condition (16) is exactly the sufficient optimality condition for problem (P_α) since $\sum_{q=1}^{Q} \alpha_q F_q(x)$ is convex. Thus, $\bar{x}$ solves (P_α).

The sufficiency follows at once since if $\bar{x}$ solves some (P_α) and there were a feasible $\tilde{x}$ for (P_v) such that $[F_1(\tilde{x}),..., F_Q(\tilde{x})] \leqslant [F_1(\bar{x}),..., F_Q(\bar{x})]$ with $F_q(\tilde{x}) < F_q(\bar{x})$ for some q, then $\sum_{q=1}^{Q} \alpha_q F_q(\tilde{x}) < \sum_{q=1}^{Q} \alpha_q F_q(\bar{x})$ would hold since $\alpha_q > 0$ for $q = 1,..., Q$. This contradicts the fact that $\bar{x}$ solves (P_α). Q.E.D.

Approximation problems having the form of problem (P_v) have been previously considered in the literature. Bacopoulos [1] considers the problem of approximating a given real-valued function by a unisolvent function simultaneously with respect to several weight functions. Johnson [11] considers the problem of uniformly approximating a vector-valued function. The approach developed in this paper permits treatment of these problems with additional side conditions such as interpolation and one-sidedness. Consider the following general vector-valued approximation problem:

$$\text{"minimize"}_{x,\tau}\ (\tau_1, ..., \tau_Q)$$

s.t.

$$\text{(i)}\quad -\tau_q \leqslant W_q(t) \left[\sum_{i=1}^{n} x_i \phi_i(t) - f_q(t) \right] \leqslant \tau_q \text{ all } t \in T,$$

for $q = 1,..., Q$,

$$\text{(ii)}\quad l(t) \leqslant \sum_{i=1}^{n} x_i \phi_i(t) \leqslant u(t) \text{ all } t \in T, \tag{17}$$

(iii) $\gamma_{1k} \leqslant \sum_{i=1}^{n} x_i \phi_i(\bar{t}^k) \leqslant \gamma_{2k}$,

for $k = 1, \ldots, K_0$,

(iv) $\sum_{i=1}^{n} x_i \phi_i(\bar{t}^k) = \gamma_k \quad$ for $\quad k = K_0 + 1, \ldots, K$,

where $l(t) < u(t)$ for all $t \in T$ with both $l(t)$ and $u(t)$ being continuous functions on T and $\gamma_{1k} < \gamma_{2k}$ for $k = 1, \ldots, K_0$, with each $f_q(t)$ and $W_q(t)$ also being continuous on T and $W_q(t) > 0$ for all $t \in T$.

THEOREM 10. *Assume that no $f_q(t)$ is in the span of $\{\phi_i(t)\}_{i=1}^n$ and that $\{\phi_i(t)\}_{i=1}^n$ is a Haar set on T. Then a point $(\tau_1^*, \ldots, \tau_Q^*, x_1^*, \ldots, x_n^*)$ which is feasible for* (17) *yields an optimal approximation if and only if the origin of R^n can be written as a convex combination of at most $n + Q$ points from the union of the sets*

$$X_q = \left\{ e_q(t) \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix} \middle| \, |e_q(t)| = \| e_q(t) \|_T = \tau_q^* \right\},$$

where $e_q(t) = W_q(t)[\sum_{i=1}^n x_i^* \phi_i(t) - f_q(t)]$ and $\| e_q(t) \|_T = \max_{t \in T} |e(t)|$,

$$X_l = \left\{ - \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i(t) = l(t) \right\},$$

$$X_u = \left\{ + \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i(t) = u(t) \right\},$$

$$X_{1k} = \left\{ - \begin{pmatrix} \phi_1(\bar{t}^k) \\ \vdots \\ \phi_n(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i(\bar{t}^k) = \gamma_{1k} \right\},$$

$$X_{2k} = \left\{ + \begin{pmatrix} \phi_1(\bar{t}^k) \\ \vdots \\ \phi_n(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i \phi_i(\bar{t}^k) = \gamma_{2k} \right\}.$$

$$X_k = \left\{ \pm \begin{pmatrix} \phi_1(\bar{t}^k) \\ \vdots \\ \phi_n(\bar{t}^k) \end{pmatrix} \middle| \sum_{i=1}^n x_i^* \phi_i(\bar{t}^k) = \gamma_k \right\},$$

with at least one point from each X_q for $q = 1, \ldots, Q$.

Proof. Problem (17) satisfies vector constraint qualification 2 because

$\{\phi_i(t)\}_{i=1}^n$ is a Haar set. This follows from the fact that by appropriately choosing parameters $\{x_i\}_{i=1}^n$, $\sum_{i=1}^n x_i\phi_i(t)$ interpolates any n values at any n distinct points t^k in T because the determinant

$$\begin{vmatrix} \phi_1(t^1) & \cdots & \phi_n(t^1) \\ & \vdots & \\ \phi_1(t^n) & \cdots & \phi_n(t^n) \end{vmatrix},$$

is nonzero by the definition of a Haar set. Thus, by setting $\tilde{y} = (0,\dots, 0, \tilde{x}_1, \dots, \tilde{x}_n)$, where $\{\tilde{x}_i\}_{i=1}^n$ interpolates the required values at the required points, $\tilde{y}$ is a vector which satisfies vector constraint qualification 2.

If $(\tau_1^*,\dots, \tau_Q^*, x_1^*,\dots, x_n^*)$ solves problem (17), then by Theorem 9, the same point solves problem (P_α) for some $\alpha \in R^Q$ with each $\alpha_q > 0$. Moreover, since (17) satisfies vector constraint qualification 2, the associated problem (P_α) satisfies constraint qualification 2. By Theorem 4, $(\tau_1^*,\dots,\tau_Q^*, x_1^*,\dots,x_n^*)$ solves (P_α) if and only if there exists integers $0 \leqslant r_0 \leqslant r_1 \leqslant r_2 \leqslant r_3 \leqslant n + Q$ together with $t^r \in T$ for $r = 1,\dots, r_1$, $1 \leqslant k_r \leqslant K_0$ for $r = r_1 + 1,\dots, r_2$, $K_0 + 1 \leqslant k_r \leqslant K$ for $r = r_2 + 1,\dots, r_3$ and $1 \leqslant q_r \leqslant Q$ for $r = 1,\dots, r_0$ such that

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_Q \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix} + \sum_{r=1}^{r_0} \lambda_r \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \\ \hline (-1)^{\epsilon_r} W_{q_r}(t^r)\,\phi_1(t^r) \\ \vdots \\ (-1)^{\epsilon_r} W_{q_r}(t^r)\,\phi_n(t^r) \end{bmatrix} + \sum_{r=r_0+1}^{r_1} \lambda_r \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hline (-1)^{\epsilon_r}\phi_1(t^r) \\ \vdots \\ (-1)^{\epsilon_r}\phi_n(t^r) \end{bmatrix}$$

$$+ \sum_{r=r_1+1}^{r_2} \lambda_r \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hline (-1)^{\epsilon_r}\phi_1(\bar{t}^{k_r}) \\ \vdots \\ (-1)^{\epsilon_r}\phi_n(\bar{t}^{k_r}) \end{bmatrix} + \sum_{r=r_2+1}^{r_3} \lambda_r \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hline (-1)^{\epsilon_r}\phi_1(\bar{t}^{k_r}) \\ \vdots \\ (-1)^{\epsilon_r}\phi_n(\bar{t}^{k_r}) \end{bmatrix} = 0 \qquad (18)$$

where $\lambda_r > 0$ for $r = 1,\dots, r_3$, $\epsilon_r = 0$ or 1 such that

(i) for $r = 1,\dots, r_0$,

$$\epsilon_r = \begin{cases} 0 \text{ if } W_{q_r}(t^r)\left[\sum_{i=1}^n x_i^*\phi_i(t^r) - f_{q_r}(t^r)\right] = \tau_{q_r}^*, \\ 1 \text{ if } W_{q_r}(t^r)\left[\sum_{i=1}^n x_i^*\phi_i(t^r) - f_{q_r}(t^r)\right] = -\tau_{q_r}^*, \end{cases}$$

(ii) for $r = r_0 + 1,..., r_1$,

$$\epsilon_r = \begin{cases} 0 \text{ if } \sum_{i=1}^n x_i^*\phi_i(t^r) = u(t^r), \\ 1 \text{ if } \sum_{i=1}^n x_i^*\phi_i(t^r) = l(t^r), \end{cases}$$

(iii) for $r = r_1 + 1,..., r_2$,

$$\epsilon_r = \begin{cases} 0 \text{ if } \sum_{i=1}^n x_i^*\phi_i(\bar{t}^{k_r}) = \gamma_{2k_r}, \\ 1 \text{ if } \sum_{i=1}^n x_i^*\phi_i(\bar{t}^{k_r}) = \gamma_{1k_r}, \end{cases}$$

and

(iv) for $r = r_2 + 1,..., r_3$, the ϵ_r is chosen appropriately so that $\lambda_r > 0$.

Since no $f_q(t)$ is in the span of $\{\phi_i(t)\}_{i=1}^n$, $\tau_q^* > 0$ for $q = 1,..., Q$. Thus $\lambda_r = \lambda_r/\tau_q^* \cdot \tau_q^*$ for $r = 1,..., r_0$ and

$$e_{q_r}(t^r) = W_{q_r}(t^r)\left[\sum_{i=1}^n x_i^*\phi_i(t^r) - f_{q_r}(t^r)\right] = \begin{cases} \tau_{q_r}^* & \text{if } \epsilon_r = 0, \\ -\tau_{q_r}^* & \text{if } \epsilon_r = 1, \end{cases}$$

and this rth constraint corresponds to the q_rth constraint of (i) being active. Normalizing the resulting coefficients such that they sum to 1, the desired result is shown. Note that there must be at least one vector from each X_q because of the first Q components of the Eq. (18) and each $\alpha_q > 0$.

Since all the steps are reversible to obtain (18), by Theorem 9 the convex combination conditions are also sufficient. Q.E.D.

Given a set of parameters $\{x_i^*\}_{i=1}^n$, a point $t^0 \in T$ is called a *positive vector-extremum* of problem (17) if for some $q = 1,..., Q$,

$$W_q(t^0)\left[\sum_{i=1}^n x_i^*\phi_i(t^0) - f_q(t^0)\right] = +\left\| W_q(t)\left[\sum_{i=1}^n x_i^*\phi_i(t) - f_q(t)\right]\right\|_T$$

and similarly t^0 is called a *negative vector-extremum* of problem (17) if for some $q = 1,..., Q$,

$$W_q(t^0)\Big[\sum_{i=1}^{n} x_i{}^*\phi_i(t^0) - f_q(t^0)\Big] = -\Big\| W_q(t)\Big[\sum_{i=1}^{n} x_i{}^*\phi_i(t) - f_q(t)\Big]\Big\|_T .$$

Thus, the vectors composing the sets X_q in Theorem 10 are evaluated at either positive or negative vector-extremums. Moreover, the theorem states that there must be at least one vector evaluated at a vector-extremum for each $q = 1,..., Q$ in the convex combination. There are said to be $n + 1$ *vector-alternates* on T for problem (17) if there are $n + 1$ $t^k \in T$ with $t^1 \leqslant \cdots \leqslant t^{n+1}$ such that the points are alternately positive and negative vector-extremum.

If constraints (ii), (iii), and (iv) in problem (17) are dropped, then it follows from Theorem 10 that $(\tau_1{}^*,..., \tau_Q{}^*, x_1{}^*,..., x_n{}^*)$ solves problem (17) if and only if there are at least $n + 1$ vector-alternates for (17). If the point solves (17), then by the Caratheodory theorem [3], the convex combination can be reduced to at most $n + 1$ points, and by the alternation lemma for Haar sets [3, p. 74], these are the $n + 1$ vector-alternates. Conversely, if there are $n + 1$ vector-alternates, and if there is no vector-extremum for some q, $1 \leqslant q \leqslant Q$, included in this convex combination, it can be inserted in the convex combination by adding some appropriate convex combination to the original one since $n + 1$ such vectors are linearly dependent. Thus, the result is a convex combination of at most $n + Q$ vectors equal to 0 with at least one vector from each set X_q. By Theorem 10 this is sufficient for $(\tau_1{}^*,..., \tau_Q{}^*, x_1{}^*,..., x_n{}^*)$ to solve (17). Consequently, Theorem 10 is a generalization of the characterization theorem developed by Bacopoulos [1].

5. Chebyshev Approximation with Nonstandard Norms

Previous sections of this paper have described characterization theorems for Chebyshev approximation problems with the standard objective of minimizing the maximum error or the vector version of the same objective. This section briefly discusses some approximation problems which have nonstandard objective functions, but which are closely related to the Chebyshev-problems. No proofs are given since they follow the general pattern used in proofs of previous sections.

First, consider the problem of Chebyshev approximation of both a function and its derivatives, as first considered by Moursund [18] and Moursund and Stroud [19].

The problem is

$$\underset{x}{\text{minimize}} \underset{\substack{k=0,1,\dots,r\\ t\in T}}{\text{maximum}} \left\{ W_0(t)\left|\sum_{i=1}^{n} x_i\phi_i(t) - f(t)\right|, \dots, W_r(t)\left|\sum_{i=1}^{n} x_i\phi_i^{(r)}(t) - f^{(r)}(t)\right|\right\},$$

where each $W_k(t) > 0$ for all $t \in T$ and both $\{\phi_i(t)\}_{i=1}^{n}$ and $f(t)$ have continuous rth derivatives for some $r > 0$. This can be rewritten as the following optimization problem:

$$\underset{\tau,x}{\text{minimize}}\ \tau$$

s.t.

$$-\tau \leqslant W_k(t)\left(\sum_{i=1}^{n} x_i\phi_i^{(k)}(t) - f^{(k)}(t)\right) \leqslant \tau \quad \text{all} \quad t \in T, \tag{19}$$

for $k = 0, 1, \dots, r$.

The solutions of (19) are characterized by the following theorem.

THEOREM 11. *A feasible point* $(\tau^*, x_1^*, \dots, x_n^*)$ *for* (19) *yields an optimal approximation for the problem if and only if the origin of* R^n *can be written as a convex combination of at most* $n + 1$ *points from the* $r + 1$ *sets*

$$X_k = \left\{ e_k(t)\begin{pmatrix} \phi_1^{(k)}(t)\\ \vdots \\ \phi_n^{(k)}(t)\end{pmatrix} \,\middle|\, W_k(t)\left|\sum_{i=1}^{n} x_i^*\phi_i^{(k)}(t) - f^{(k)}(t)\right| = \tau^* \right\},$$

where $e_k(t) = W_k(t)[\sum_{i=1}^{n} x_i^*\phi^{(k)}(t) - f^{(k)}(t)]$ *for* $k = 0, 1, \dots, r$.

Next, consider the problem of Chebyshev approximation of a function and its derivatives as developed by Laurent [15]. A generalized version of this problem is

$$\underset{x}{\text{minimize}} \underset{t\in T}{\text{maximum}} \left\{ \sum_{k=0}^{r} W_k(t)\left|\sum_{i=1}^{n} x_i\phi_i^{(k)}(t) - f^{(k)}(t)\right|\right\},$$

where both $\{\phi_i(t)\}_{i=1}^{n}$ and $f(t)$ have continuous rth derivatives and each $W_k(t) > 0$ for all $t \in T$. This can be rewritten as an optimization problem:

$$\underset{x,\tau}{\text{minimize}} \sum_{k=0}^{r} \tau_k$$

s.t.

$$-\tau_k \leqslant W_k(t)\left[\sum_{i=1}^{n} x_i\phi_i^{(k)}(t) - f^{(k)}(t)\right] \leqslant \tau_k \qquad \text{for} \quad k = 0,\dots, r. \tag{20}$$

The solutions of (20) are characterized by the following theorem.

THEOREM 12. *A feasible point* $(\tau_1^*,..., \tau_r^*, x_1^*,..., x_n^*)$ *for* (20) *yields an optimal approximation for the problem if and only if the origin of* R^n *can be written as a convex combination of at most* $n + 1$ *points from the sets*

$$X_k = \left\{ e_k(t)\begin{pmatrix} \phi_1^{(k)}(t) \\ \vdots \\ \phi_n^{(k)}(t) \end{pmatrix} \middle| \, W_k(t)\left| \sum_{i=1}^{n} x_i^* \phi_i^{(k)}(t) - f^{(k)}(t) \right| = \tau_k^* \right\},$$

where $e_k(t) = W_k(t)(\sum_{i=1}^{n} x_i \phi^{(k)}(t) - f^{(k)}(t))$, *with at least one point from each set* X_k *for* $k = 0,..., r$.

Finally, it should be observed that additional problems such as relative error approximation with the objective

$$\underset{x}{\text{minimize}}\ \underset{t \in T}{\text{maximum}}\ \frac{|\sum_{i=1}^{n} x_i \phi_i(t) - f(t)|}{|f(t)|}$$

can easily be handled by the techniques of this paper. Also, it should be obvious that solutions to each of the problems posed in this section could be characterized when additional constraints such as interpolation, one-sidedness, and monotonicity are present. This is not done here for the sake of brevity.

6. CONSTRAINED L^1 APPROXIMATION

The problems treated in the previous sections have all been concerned with L^∞ approximation. The purpose of this section is to develop similar characterization theorems for solutions of problems of L^1 approximation. Although the results are not developed in the fullest possible generality, the theorems proved here are sufficient to illustrate the potential of this approach to developing characterization theorems for general L^1 approximation problems.

Throughout this section it shall be assumed that $T = [a, b]$, a closed bounded interval of the real line. Consider the following general L^1 approximation problem

$$\underset{x}{\text{minimize}} \int_a^b \left| \sum_{i=1}^{n} x_i \phi_i(t) - f(t) \right| dt$$

s.t.

(i) $l_k(t) \leqslant \sum_{i=1}^{n} x_i \phi^{(j_k)}(t) \leqslant u_k(t), \quad \text{all} \quad t \in [a, b],$

$$\text{for} \quad k = 1, \ldots, K_0, \tag{21}$$

$$\text{(ii)} \quad \gamma_{1k} \leqslant \sum_{i=1}^{n} x_i \phi_i^{(j_k)}(\bar{t}^k) \leqslant \gamma_{2k},$$

$$\text{for} \quad k = K_0 + 1, \ldots, K_1.$$

$$\text{(iii)} \quad \sum_{i=1}^{n} x_i \phi_i^{(j_k)}(\bar{t}^k) = \gamma_k,$$

$$\text{for } k = K_1 + 1, \ldots, K,$$

where the indices j_k are chosen nonnegative integers, for each $k = 1, \ldots, K_0$, $l_k(t) \leqslant u_k(t)$ for all $t \in [a, b]$ with both $l_k(t)$ and $u_k(t)$ being continuous functions on $[a, b]$, each $\bar{t}^k \in [a, b]$, and $\gamma_{1k} < \gamma_{2k}$ for $k = K_0 + 1, \ldots, K_1$.

In order to apply the theorems of Section 1 to problem (21), the following lemma concerning differentiability is needed. Define the function

$$F(x) = \int_a^b \left| \sum_{i=1}^{n} x_i \phi_i(t) - f(t) \right| dt.$$

LEMMA 1. *If $\sum_{i=1}^{n} x_i \phi_i(t) - f(t)$ has only a finite number of zeros in $[a, b]$, then $F(x)$ is continuously differentiable with*

$$\frac{\partial F(x)}{\partial x_i} = - \int_a^b \phi_i(t) \operatorname{sgn} \left(f(t) - \sum_{i=1}^{n} x_i \phi_i(t) \right) dt,$$

for $i = 1, \ldots, n$, where the function sgn *is defined by*

$$\operatorname{sgn}(g(x)) = \begin{cases} +1 & \textit{if} \quad g(x) > 0, \\ 0 & \textit{if} \quad g(x) = 0 \\ -1 & \textit{if} \quad g(x) < 0. \end{cases}$$

Proof. Let $t_1, \ldots, t_k$ be the roots of $\sum_{i=1}^{n} x_i \phi_i(t) - f(t)$ in $[a, b]$. For sufficiently small $\epsilon > 0$, define $A = [a + \epsilon, t_1 - \epsilon] \cup [t_1 + \epsilon, t_2 - \epsilon] \cup \cdots \cup [t_k + \epsilon, b - \epsilon]$ and $B = [a, b] \cap A^c$. Define

$$\delta = \min \left\{ \left| f(t) - \sum_{i=1}^{n} x_i \phi_i(t) \right| t \in A \right\} > 0,$$

and observe that if $0 < \| \lambda \phi_i(t) \| < \delta$, $[a, b]$ then,

$$\operatorname{sgn} \left(f(t) - \sum_{i=1}^{n} x_i \phi_i(t) - \lambda \phi_i(t) \right) = \operatorname{sgn} \left(f(t) - \sum_{i=1}^{n} x_i \phi_i(t) \right).$$

on A. Then for small enough λ (either positive or negative) it can be shown that

$$\int_B \phi_i(t)\operatorname{sgn}\left(f(t)-\sum_{i=1}^n x_i\phi_i(t)\right)dt - \int_B |\phi_i(t)|\,dt$$

$$\leqslant \frac{F(x+\lambda e_i)-F(x)}{\lambda} + \int_a^b \phi_i(t)\operatorname{sgn}\left(f(t)-\sum_{i=1}^n x_i\phi_i(t)\right)dt$$

$$\leqslant \int_B \phi_i \operatorname{sgn}\left(f(t)-\sum_{i=1}^n x_i\phi_i(t)\right)dt + \int_B |\phi_i(t)|\,dt,$$

where e_i is the unit vector in the ith component. Observe that the left-hand side of this inequality is bounded below by $-4k\epsilon \|\phi_i(t)\|_{[a,b]}$ and that the right-hand side is bounded by $4k\epsilon \|\phi_i(t)\|_{[a,b]}$. By choosing λ sufficiently small, $\epsilon > 0$ can be made arbitrarily small, which proves that

$$\frac{\partial F(x)}{\partial x_i} = \lim_{\lambda\to 0}\frac{F(x+\lambda e_i)-F(x)}{\lambda} = -\int_a^b \phi_i(t)\operatorname{sgn}\left(f(t)-\sum_{i=1}^n x_i\phi_i(t)\right)dt,$$

as desired. Q.E.D.

The following theorem characterizes the solutions of problem (21).

THEOREM 13. *Assume that problem* (21) *satisfies either constraint qualification* 1 *or* 2. *Then a point* $(x_1^*,\ldots,x_n^*)$, *feasible for problem* (21) *with* $\sum_{i=1}^n x_i^*\phi_i(t) - f(t)$ *having only a finite number of zeros, solves* (21) *if and only if there are integers* $0 \leqslant s_0 \leqslant s_1 \leqslant s \leqslant n$ *together with* $t^q \in [a,b]$ *for* $q = 1,\ldots,s_0$, $1 \leqslant k_q \leqslant K_0$ *for* $q = 1,\ldots,s_0$, $K_0+1 \leqslant k_q \leqslant K_1$ *for* $q = s_0+1,\ldots,s_1$ *and* $K_1+1 \leqslant k_q \leqslant K$ *for* $q = s_1+1,\ldots,s$ *and real numbers* $\lambda_q \neq 0$ *such that*

$$\int_a^b P(x;t)\operatorname{sgn}\left(f(t)-\sum_{i=1}^n x_i^*\phi_i(t)\right)dt$$

$$= \sum_{q=1}^{s_0}\lambda_q P^{(j_{k_q})}(x;t^q) + \sum_{q=s_0+1}^{s}\lambda_q P^{(j_{k_q})}(x;\bar t^{k_q}) \tag{22}$$

for all generalized polynomials $P(x;t) = \sum_{i=1}^n x_i\phi_i(t)$ *where* $P^{(j)}(x;t) = \sum_{i=1}^n x_i\phi^{(j)}(t)$ *and where the sign of* λ_q *is determined by*

(i) *for* $q = 1,\ldots s_0$,

$$\lambda_q > 0 \text{ if} \sum_{i=1}^n x_i * \phi_i^{(j_{k_q})}(t^q) = u_k(t^q),$$

$$\lambda_q < 0 \; if \sum_{i=1}^{n} x_i{}^* \phi_i^{(j_{k_q})}(t^q) = l_{k_q}(t),$$

(ii) *for* $q = s_0 + 1,..., s_1$,

$$\lambda_q > 0 \; if \sum_{i=1}^{n} x_i{}^* \phi_i^{(j_{k_q})}(\bar{t}^{k_q}) = \gamma_{2k_q},$$

$$\lambda_q < 0 \; if \sum_{i=1} x_i{}^* \phi_i^{(j_{k_q})}(\bar{t}^{k_q}) = \gamma_{1k_q},$$

and the signs for λ_q *for* $q = s_1 + 1,..., s$ *are indeterminant.*

Proof. The fact that the objective function $F(x)$ is convex follows from the triangle inequality. Thus, by Theorem 4 and Lemma 1, $(x_1{}^*,..., x_n{}^*)$ solves (21) if and only if

$$-\int_a^b \phi_i(t) \operatorname{sgn} \Big(f(t) - \sum_{i=1}^{n} x_i^* \phi_i(t)\Big) \, dt + \sum_{q=1}^{s_0} \lambda_q \phi_i^{(j_{k_q})}(t^q)$$

$$+ \sum_{q=s_0+1}^{s} \lambda_q \phi_i^{(j_{k_q})}(\bar{t}^{k_q}) = 0, \tag{23}$$

holds for each $i = 1,..., n$ with the integers s_i and parameters λ_q defined as in the theorem and

$$t^q \in \Big\{ t \in [a, b] \Big| \sum_{i=1}^{n} x_i{}^* \phi_i^{(j_{k_q})}(t) = u_k(t) \text{ or } l_k(t) \Big\},$$

for $q = 1,..., s_0$,

$$\bar{t}^{k_q} \in \Big\{ \bar{t}^k \in [a, b] \Big| \sum_{i=1}^{n} x_i{}^* \phi_i^{(j_{k_q})}(\bar{t}^k) = \gamma_{1k_q} \text{ or } \gamma_{2k_q} \Big\},$$

for $q = s_0 + 1,..., s_1$.

The conclusion follows directly from (23). Q.E.D.

Constraint qualification 1 holds for problem (21) if $\{\phi_i(t)\}_{i=1}^n$ is a Haar set on $[a, b]$, there are no derivatives in the constraints of (21), i.e., $j_k = 0$ for $k = 1,..., K$, and $l_0(t) < u_0(t)$ for all $t \in [a, b]$. Thus, Theorem 13 applies immediately to a wide variety of problems without being concerned whether or not a constraint qualification is satisfied. Furthermore, for one-sided approximation, say $f(t) \geqslant \sum_{i=1}^n x_i \phi_i(t)$, Theorem 13 can be derived without any condition on the roots of $\sum_{i=1}^n x_i \phi_i(t) - f(t)$. Thus, these results generalize the characterization theorem previously developed by de Vore [6].

We note that general L^p approximation problems with $1 < p < \infty$ can be handled in exactly the same manner as the L^1 problem. Furthermore, much more general forms of Lemma 1 can be given which would be less restrictive in Theorem 13. However, the purpose of this paper has been to explain the basic types of problems which can be treated by this approach rather than the most general in each case.

7. Conclusions

The underlying theme throughout this paper has been that characterization theorems for solutions of a wide variety of L^p approximation problems can be obtained in a simple and unified manner by using a mathematical optimization approach. In addition to the unity it lends to the development of characterization theorems, the mathematical programming approach is well-suited for (i) development of efficient algorithms for obtaining best approximations by using algorithms which solve the associated mathematical programming problems, and (ii) development of error estimates for an approximation problem by using the dual optimization problem which is always associated with the original optimization formulation of the approximation problem. Future papers will explore both of these aspects. Of particular interest is an algorithm, closely related to the second algorithm of Remez, which solves general optimization problems of the form (P) described in Section 1 [9].

References

1. A. Bacopoulos, Nonlinear Chebyshev approximation by vector-norms, *J. Approximation Theory* **2** (1969), 79–84.
2. R. Bojanic and R. De Vore, On polynomials of best one sided approximation, *L'Enseignement Math.* **12** (1966), 139–144.
3. E. W. Cheney, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
4. F. Deutsch, On uniform approximation with interpolatory constraints, *J. Math. Anal. Appl.* **24** (1968), 62–79.
5. F. R. Deutsch and P. H. Maserick, Applications of the Hahn-Banach theorem in approximation theory, *SIAM Rev.* **9** (1967), 516–530.
6. R. De Vore, One-sided approximation of functions, *J. Approximation Theory* **1** (1968), 11–25.
7. K. R. Gehner, Optimization problems with an infinite number of constraints and applications to constrained approximation problems, Ph.D. thesis, University of Wisconsin, Madison, WI, 1971.
8. K. R. Gehner, Necessary and sufficient optimality conditions for the Fritz John problem with linear equality constraints, *SIAM J. Control* **12** (1974), 140–149.
9. K. R. Gehner, An algorithm for solving an infinitely-constrained mathematical program, *SIAM J. Numer. Anal.*, to be submitted.

10. F. JOHN, Extremum problems with inequalities as subsidiary conditions, *in* "Studies and Essays, Courant Anniversary Volume," pp. 187–204, Interscience, New York, 1948.
11. L. W. JOHNSON, Uniform approximation of vector-valued functions, *Numer. Math.* **13** (1969), 238–244.
12. R. C. JONES AND L. A. KARLOVITZ, Iterative construction of constrained Chebyshev approximation of continuous functions, *SIAM J. Numer. Anal.* **5** (1968), 574–585.
13. J. W. KAMMERER, Optimal approximation of functions: one-sided approximations and extrema preserving approximations, Ph.D. thesis, University of Wisconsin, Madison, WI, 1959.
14. P. J. LAURENT, Approximation uniforme de fonctions continues sur un compact avec constraintes de type inegalitie, *R.I.R.O.* **1** (1967), 81–95.
15. P. J. LAURENT, Théorèmes de caractérisation d'une meilleure approximation dans un espace norme et généralisation de l'algorithme de Remes, *Numer. Math.* **10** (1967), 190–208.
16. G. G. LORENTZ AND K. L. ZELLER, Monotone approximation by algebraic polynomials, *Trans. Amer. Math. Soc.* **149** (1970), 1–18.
17. O. L. MANGASARIAN, "Nonlinear Programming," McGraw-Hill, New York, 1969.
18. D. G. MOURSUND, Chebyshev approximation of a function and its derivatives. *Math. Comp.*, **18** (1964), 382–389.
19. D. G. MOURSUND AND A. H. STROUD, The best Cebysev approximation to a function and its derivatives on $n+2$ points, *SIAM J. Numer. Anal.* **2** (1964), 15–23.
20. J. R. RICE, Approximation with convex constraints, *SIAM J. Appl. Math.* **11** (1963), 15–32.
21. T. J. RIVLIN AND H. S. SHAPIRO, A unified approach to certain problems of approximation and minimization, *SIAM J. Appl. Math.* **9** (1961), 670–699.
22. L. L. SCHUMAKER AND G. D. TAYLOR, On approximation by polynomials having restricted ranges II. *SIAM J. Numer. Anal.*, **6** (1969), 31–36.
23. G. D. TAYLOR, On approximation by polynomials having restricted ranges, *SIAM J. Numer. Anal.*, **5** (1968), 258–268.
24. G. D. TAYLOR, Approximation by functions having restricted ranges III. *J. Math. Anal. Appl.*, **27** (1969), 241–248.
25. G. D. TAYLOR, Approximation by functions having restricted ranges: equality case. *Numer. Math.*, **14** (1969), 71–78.
26. G. D. TAYLOR AND M. J. WINTER, Calculation of best restricted approximations. *SIAM J. Numer. Anal.* **7** (1970), 248–255.